International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Short- and Long-Term Effective Time Series Stock Market Cost Predication Methods Using Learning Techniques

Sourabh Jain, Dr. Navdeep Kaur Saluja, Dr. Anil Pimplapure

Submitted: 06/02/2024 **Revised**: 14/03/2024 **Accepted**: 22/03/2024

Abstract: The Stock market cost is an excellent example of the time series data. The trust worthy investment options is essential for increasing the successful stick investment enhancement. Many times, investors are hesitant to decide the best stock prizes. Therefore, using machine learning (ML) might predict the stock prizes in advance. Paper presented the validation of the ML based stock market prediction for the Microsoft database. This paper considered the stock market prediction (SMP) problem as the solution via time series prediction. The continually varying nature of stock market makes it non stationary series of data this makes the cost prediction a tough problem. To increase prediction accuracy this paper proposed to examine time series auto regressive models based on optimal degree of differencing's. The lag order is varied for achieving best production and degree of differencing's to make model nearly stationary. The prediction modes ARIMA, SARMAX and LSTM are trained on the large Microsoft stock dataset for last 20 years. For short term prediction the ARIMA is proposed and for the long term 150 days prediction the LSTM is found to be best. Then based on the proposed model the stock cost prediction is tested using the regression model on the recent Microsoft data. The results are evaluated based on the mean square error (MSE) and absolute difference error (ADE) values for different models. Significant SMP accuracy is achieved for 150 days around 93% over rich data sample sizes.

Keywords*: Stock Market Prediction, Machine Learning, Time Series Prediction, ARIMA, SARMAX, LSTM, closing cost, Opening Cost, Training, Testing MSE.*

1 Introduction

The trustworthy stock cost production is most relevant and essential field of the research in recent scenario. With the increasing investment market worldwide, it becomes essential to design simple and effective stock cost prediction approaches. When predicting the stock market data, one must make the assumption that any information that is now open to the public has a certain predictive value for the future of stock returns. The stock market is a tremendously difficult area for fascinating industry.

The basic component of underlying stock price estimates is the projected stock prices in the near future. This paper attempts to forecast the relatively short-term target values for such stocks. To enhance market prices as well as investment planning, the SMP technique was adopted. There are many SMP et al., 2015) and hGWO-PS/2DOF-PID (Soni, V., et al., 2016a, 2016b, 2017, 2020). approaches are available in the literature every algorithm attempts to improve the prediction accuracy. The broad classification of the SMP methodologies is given in the Figure 1. The most frequently used machine learning (ML) approaches includes neural network [1], linear regression [2], support vector machine (SVM) [3], soft computing [4], Three are various time series prediction methodologies as. Method of prediction using Long-Short-term Memory LSTM [5] van predict long term data, Autoregressive integrated moving average (ARIMA) [6, 7 and 8], SARIMA [9 and 10] (fits for short term prediction), deep learning-based prediction [11, 12 and 13] only limitations that it is used for the long-term production and need lager dataset.

Table 1 Abbreviations used in manuscript.

Abbreviation	Description	Abbreviation	Description
ML	Machine Learning	MSE	Mean square error
SMP	Stock Market Prediction	TS	Time Series
ARIMA	Autoregressive Integrated Moving Average	RMSE	Root of MSE
AR	Autoregressive	MA	Moving Average

[1]*Research Scholar, Department of Computer Science and Engineering, Eklavya University Damoh Email: sourabhjain1788@gmail.com*
[2]*Professor, Department of Computer Science and Engineering, Eklavya University Damoh Email: Navdeep.saluja2000@gmail.com*
[3]*Professor Department of Computer Science and Engineering, Eklavya University Damoh Email: pimpalapureanil@gmail.com*

LSTM	Long-Short-term Memory	MAE	Mean Absolute Error
SVM	Support Vector Machine	ADE	Absolute Difference Error
RNN	Recurrent Neural Network	VAR	Vectors Auto-Regression
NSE	National Stock Exchange (NIFTY)	SARIMAX	seasonal autoregressive combined moving average with endogenous

Since stock data is has grate time series properties 'therefore this paper aimed to investigate and review states of art stock prediction methodologies especially time series prediction.

Despite extensive study on the financial markets, no universal solution for assessing or forecasting the markets have been developed and it's still an open challenge. Time series (TS) is suitable for reflecting the whole stock market and forecasting its short-term trend. Ample amount of TS techniques has been used by technical and quantitative methods to try and predict the price. However, there hasn't been much study on the suitable short-time length of data needed to anticipate stock markets. The major challenges of the SMP problem are shown in the Figure 1. Accuracy of the prediction highly depends on the quality of the data and size of sample. Rich set of data is required even for short term prediction. The major challenge in stock prediction is the occlusion or missing data consistency.

Fig 1 Major challenges of stock (SMP) production problem

2 Classification of ML based SMP Methodologies

These days, a variety of factors, including firm characteristics and associated news, political developments, natural calamities, etc, have an impact on stock prices

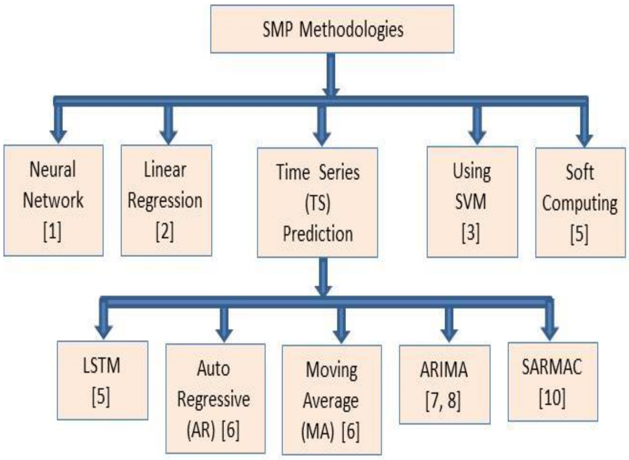


Fig 2 Classification of various ML based SMP methodologies

. An extensive data expected research taken care in last decade for Indian stock market values is the prime motivation behind the recent work. The broad classification of the various ML based SMP methodologies are given in the Figure 2. The exponential growth in the market values is causing requirement of research in this field and this is the reason to exponential increase in research for stock prediction. Each of these categories of the SMP is individually reviewed in the next section with special focus on autoregressive integrated moving average (ARIMA) short term model and LSTM for Long term prediction.

3 Review of Stock Prediction

Pang, X., et al. [1] have proposed a novel neural network (NN) strategy for improving market forecasts. They used live information from marketplace and illustrate the usage of Audiovisual Objects for analyzing stocks. The findings of study indicate the deeper the LSTM without integrated layer performs better. Vaishnavi Gururaj et al. [2] have proposed to design the combination of Regression with linearity (LR) lowest-squares supported vector machine (LS-SVM) as SVM variant for the stock cost prediction. It is stated that independent algorithm has less efficiency thus it required to combine other algorithm together to improve the prediction. Cao, R. et al. [3] have evaluated Chinese stock exchange data using text sentiment estimation method for Internet financial details. When paired with stock cost data, support vector machines are implanted to analyses and predict the share price. Göçken, M. et al. [4] have proposed hybrid soft competing approach taking advantage of Regressive Tree and forecasting of stock prices in this study using Unity Search. The parametric tuning is achieved using soft computing for RNN. In order to improve the predicted accuracy of predicting structure, Sidra Mehtab et al. [5] have developed deep learning enabled regression techniques utilising LSTM systems and a novel technique known as walk-forward verification. They made use of information from the NIFTY 50 index, which stands in for the National Stock Exchange (NSE) of India.

A. Review of Short-Term TS Prediction

This section considered the related works of Time series prediction. Kamalakannan J et al [6] has forecasted the behaviour of stock industry of Indian market. They have estimate and visualize results using the TS prediction approach. A mathematical ARIMA model is fitted on historical data. Because the Autoregressive Integrated Moving Average, or ARIMA, model is reliable, effective, and possesses a great deal of application for a short-term market shares foresight, it has been widely employed in the fields of finances and economics. Support vector regression (SVR), and vectors auto-regression (VAR), and LSTM for SMP were all evaluated by Majumder, A. et al. [7] using a variety of performance metrics, such as variance explanation score (EVS), mean square error (MSE), average absolute error (MAE), root average square error (RMSE), and R lined score (R2 Score). The researchers came to the conclusion that the LSTM model performed better than all of the other approaches. Adebiyi Ariyo Ariyo et al [8] have described a procedure for developing an ARIMA-based TS stock price prediction model. A price of a stock prediction model is constructed and deployed with available and the NASDAQ Stock Exchange (NYSE). The collected results showed that the ARIMA approach could compete effectively with current methods for price forecasting and has a significant promise for immediate forecasting. Daryl et al [9] have used the "SARIMA" algorithm during our study to forecast the company's stock values. The "SARIMA" approach is a traditional model that forecasts the stock market using data. This is due to the fact that the prices of stocks fluctuate over a period of time and aren't static, as "SARIMA" is able to forecast. But accuracy is untenable for long sequence forecasting. Fahad Radhi Alharbi et al [10] presented system to predict the long-term viability of the power consumption and peak load capacities using a seasonal autoregressive combined moving average with endogenous components (SARIMAX) models. But models lack excellent precision.

Sidra Mehtab et al [11] proposed stock prices methodology that combines machine learning, deep learning (DL), and statistical approaches. They traded data of National Stock Exchange (NSE) of Indian. Nusrat Rouf et al [12] improved the precision by using ML algorithm like ensemble learning and textual analysis. Approaches take into account the study and forecasting of the market for shares performed by predictive methods. Jingyi Shen et al. [13] have used pre-processing of a stock data set, used several features engineering methods, and employed DL driven algorithm for SMP suggested solution. Method achieved an overall high degree of accuracy. S. Kulshreshtha et al [14] captured real-time S&P 500 market information using a current app programming interface (API), this paper suggests a novel hybrid approach that combines the RNN method to LSTM model for TS forecasting and compared ARIMA model results. They proposed to capture both linear and irregular parts of the time sequence using an innovative combination of LSTM-ARIMA models. They choe better stock market forecast system among both and select them for data.

Shakir Khan et al [15] used ARIMA and evaluated five years' worth of Netflix's stock past information and obtain a reliable forecasting stock model. They used ARIMA (1,1,33) which outperformed the other two approaches, demonstrating the possibility of employing the ARIMA framework for precise stock predictions.

B. **Long term SMP Review**.

Aminimehr, A et al [16] evaluated various preprocessing techniques for LSTM based SMP study. The LSTM'sis compared with random forest (RF) using the deep feature mining technique. Then, the accuracy of the models is verified by calculating the MAE, MSE, MAPE, CSP, and CDCP. But data for only six months are used in this study. Sidharth Tiwary et al [17] have forecasted the E-vehicles' stock values and the closing cost of their stocks. The ARIMA (p, d, q) has been applied to forecast. Metrics like Akaike Dissemination Criterion (AIC), Log-likelihood (LL), and Behavioural Info Criteria (BIC), besides to the fundamental prediction precision metrics, are taken into account in predictions. Jimmy Ming Tai Wu et al [18] suggested an innovative architecture which utilizes Convolution Neural Network (CNN) and LSTM to accomplish precise SMP forecast based on features of finance TS price predicting task. The name of this new technique is stock sequencing array neural LSTM (SACLSTM), quite fittingly. Ten stocks from the Taiwanese and the United States are used as the experimental information. Method required large processing or elapsed time. Kelum Gajamannage et al [19] have proposed using ANN structures & the method of training have a significant influence on the ability to make appropriate real-time SMP forecasts of time periods series. Chance Robinson et al. [20] used data of sentiment from Twitter and other economic media fed into RNN based LSTM Networks, the used data over a 12-year period, and they saw for three NASDAQ equities at the conclusion of every day of trading.

Salman Ahmed et al [21] suggested technique minimizes discrepancy among the real and predicted averages of Forex candles by integrating the Foreign Exchange Loss Functions (FLF) into a LSTM model. The Forex Loss Factor was developed as a result of numerous tests. Tawum Juvert Mbah et al [22] presented mathematical modeling and forecasting of stone price volatility carried out using ANN, ARIMA and RNN. To replicate the forecasts, the RNN makes use of layers of LSTM drop outs regularization, functions for activation, mean square errors (MSE), and the optimization algorithm Adam. They used predetermined parameters and methods of LSTM by simulating future prices data throughout time. A method of statistical analysis called the ARIMA used for short term trajectory, and calendar. Faraz Ahmad et al. [23] have considered various artificial intelligence and deep learning models. Area of research includes a variety of strong models, such as ARIMA and VAR, and the final group consists of instructional techniques that use machines training and techniques from deep learning. Hum Nath et al [24] have preseted food use of the LSTM based long term SMP forecasting. The RNN and LSTM model are trained using deep or dense NN layer architecture to predict Stock prizes. Table 2 summarizes most relevant works.

Table 2 Summary of the review work for Stock provocation

S. No.	Authors Name	Methodology	Evaluation Parameters
1.	Pang, X., et al. [1]	Used LSTM based SMP prediction and stated to design deep LSTM.	The prediction error as MSE is evaluated with learning rates.
2.	Vaishnavi Gururaj et al. [2]	Proposed LS-SVM for the regressive production of the SMP data	evaluated prediction modeling accuracy
3.	Sidra Mehtab et al. [5]	Deep learning enabled regression techniques based on LSTM systems. They used the National Stock Exchange (NSE) of India. data,	Plotted root MSE (RMSE) error, mean, min and Standard deviation (SD) for days of weeks
4.	Ariyo Ariyo et al [8]	Developed an ARIMA-based TS stock price prediction model. stock information from the Nigerian Stock Exchange (NSE)	Prediction error as parameter of evaluation
5.	Fahad Radhi Alharbi et al [10]	Presented to predict the long-term viability of power consumption and peak load capacities using a SARIMAX models.	Measures accuracy and precision But models lack excellent precision.
6..	Jingyi Shen et al. [13]	Employed DL driven algorithm for SMP suggested solution. Using RNN and LSTM	high accuracy for stock market trend.
7.	Tawum Juvert Mbah et al [22]	Used ARIMA and RNN to replicate the forecasts. The RNN makes use LSTM layer drop outs regularization, functions for activation,	Employing mean square errors (MSE), and the optimization algorithm Adam for evaluation.
8.	R. Kumar et al [22]	hybrid approach as combination of SVM-LSTM approach	Correlation between a stock price and the polarity of the news. Results demonstrate the potential of suggested method for stock price forecasting at a specific time
9.	Faraz Ahmad et al. [23]	Have employed variety of strong deep learning based models, such as ARIMA and VAR for SMP	volatility in the stock market has a far greater impact on stock volatility and cost variability.
10.	Hum Nath et al [24]	Used RMM and LSTM with multi dense layer for the long-term stock prices trend prediction.	RMSE, Correlation Coefficient (R). Mean Absolute Percentage Error (MAPE) are evaluated.
11	Proposed	Uses ARIMA for short term and LSTM for long term SMP	MSE, MAE and ADE are evaluated

4 Time Series Recreation Methodologies

Time series forecasting is a difficult problem without an easy fix. Many statistical frameworks compete with each other, yet it is seldom clear which model performs best. Our projection will be determined by a technical examination of historical Microsoft data employing an ARIMA Algorithm. The Autoregressive Integrated Moving Average, or ARIA, model continues to be actively utilized in banking and finance since it is recognized to be resilient, efficient, and has a high potential of short-term SMP.

4.1 ARIMA Model

According to observation, ARMA-based models are frequently an excellent place to start. They are suitable as a foundational model in every time series problem that can obtain respectable results on the majority of time-series problems. This article provides a thorough, approachable explanation of ARIMA-based models for researchers.

The ARIMA model acronym stands for "Auto-Regressive Integrated Moving Average" and for this article will break it down into AR and MA. In nut shale ARIMA model consists of a constant summer the product of the multipliers for the AR and MA lags, as well as white noise. This formula serves as the foundation for all subsequent models and provides a framework for numerous forecasting models in various fields. The ultimate aim of the ARIMA model is to predict the values of three lag values of d, p and q as

$$Predict_{ARIMA} = \min_{MSE/MAE} f(p, d, q) \quad (1)$$

4.2 SARIMAX Models

Although the ARIMA model performs exceptionally well, it can be extremely beneficial to integrate seasonality and exogenous variables exclusively. One must choose an alternative model because the ARIMA model relies on the assumption that the time series is steady. but in most case data is non-stationary. The SARIMAX model incorporates exogenous variables, or, to put it another way, makes use of outside data to make predictions. Exogenous variables can be found in the actual world in the form of exchange rates, gold and oil prices, and outdoor temperatures. It's intriguing to consider that the historical model projection still technically models all exogenous components indirectly. Nevertheless, the model will react to its influence far more quickly if external data is included than if we rely just on the influence of lagging components.

4.3 LSTM:

The short-term data production is achieved using the long-term-short-term (LSTM) model. It's widely being opted for stationary TS data forecasting. The LSTM is extension of recurrent neural network (RNN) and is preferred for the long-term prediction of TS data.

5 Proposed SMP Methodology

In this paper the SMP problem is proposed to solve using the short-term time series prediction (TSP) model based on the ARIMA, SARMAX, and LSTM based machine learning models. The python code using pandas is written for prediction and validation of data. The data pre-processing is used for determining the lag values of the ARIMA model fitting. The proposed SMP system block diagram is shown in the Figure 3. There are three prediction models are evaluated and tested as shown in the Table 3.

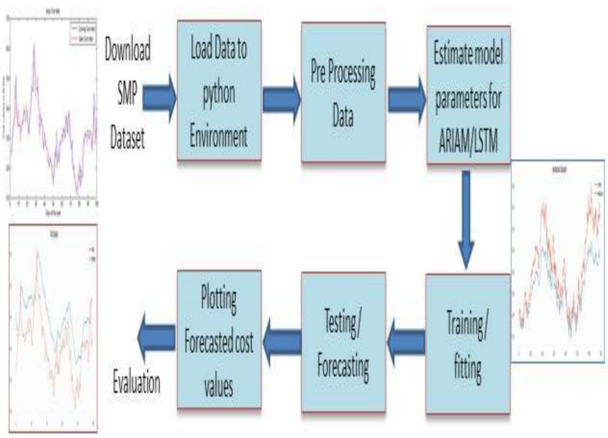


Fig 3 Proposed methodology block diagram for the SMP current research

Table 3 TS prediction environment / models used in this research

Model	Environment
ARIMA	Python 3.6.0, and Pandas, (ARIMA(2,1,2))
SARIMAX	Python 3.6.0, and Pandas,
LSTM	Python 3.6.0, and Pandas, Min-Max scalor Normalization : Multilayer LSTM Architecture

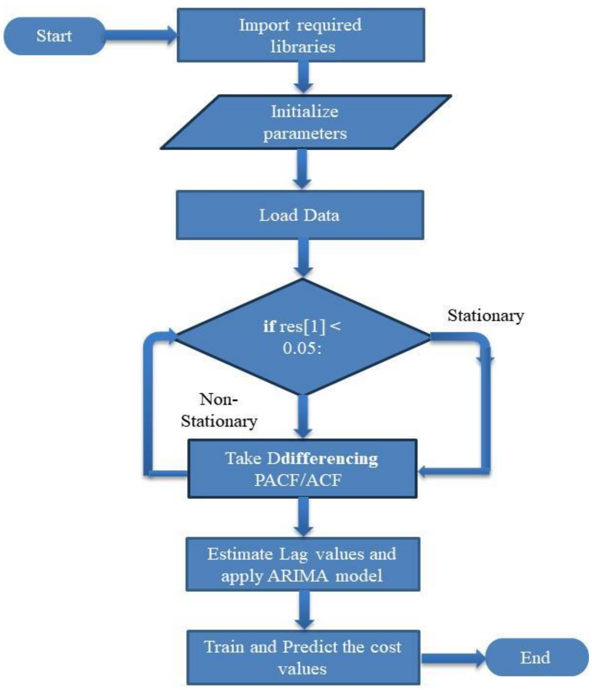
First ARIMA model is tested and since model required a stationary time series thus proposed method first check for stationarity.

5.1 ARIMA short term Model

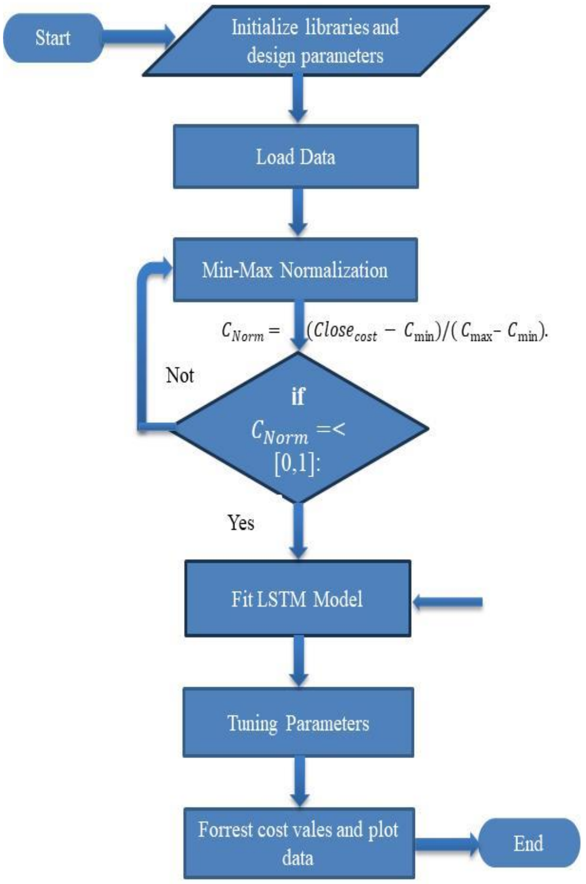
The flow chart of the proposed methodology is given in the Figure 5.a) It can be closely observed from the Figure 4 that if the data is non-stationary, proposed method computes the first-level differencing and recheck the stationarity. Given the abundance of NAN values within the raw data, thus paper proposed to fill those NAN values using historical cost values.

5.2 Long-term LSTM model

This study predicts the closing price of the stock market for Microsoft data of last 6 years. the 100 and 150 samples or the following day using a specific neural network design called LSTM model for long term case the proposed LSTM prediction system shown in Figure 6. It is proposed to normalize the data and then apply LSTM model for the long-term cost prediction the quantitative evaluation is based on the MSE and MAR error evaluation.



a) Flowchart of ARIMA model



b) Flow chart of the LSTM Model

Fig 5 systematic SMP proposed TS based cost prediction low chart

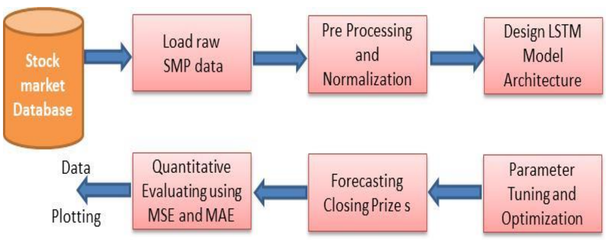


Fig 6 Proposed LSTM based TS long term SMP system

Initially data is pre-processed to feed it to LSTM model. The previous time values in a TS influence the future time series values, therefore paper created our LSTM training features using lags of the opening price of the stock. Initially the training is started with the lag value of 10. In the paper it is proposed to scale the data by employing the MinMaxScalar provide in sci-kit learn library the Min/Max is used as the normalization method at the pre-processing stage of data. Since this is a time series forecast it must not be forecasting too deep in future because the model will not perform well, and it makes no sense to predict too deep in future specially the stock prediction.

Consequently, in the event when the variable Closing cost's Min and Max values are $C_{\max}$ and $C_{\min}$ respectively, then the respective normalized value C_{norm} is measured as:

$$C_{Norm} = (Close_{cost} - C_{\min})/(C_{\max} - C_{\min}) \quad (2) .$$

It is to observe here that all values of C_{Norm} after the normalisation process fall within the range [0, 1]. Paper will predict the future data in for next 100 days and 150 says the validation data chosen is 15% of the training data. Once the basic LSTM is validated then the testing is applied using parametric optimization over testing data.

6 Results and Discussions

This paper attempts to forecast the relatively short-term target values for such stocks. To enhance market prices as well as investment planning, the SMP technique was adopted. This paper considered the stock market prediction (SMP) problem as the solution via time series prediction. Time series (TS) is suitable for reflecting the whole stock market and forecasting its short-term trend. The snapshot of the Microsoft data only after 2018 is represented in the Figure 5. It can be observed that total a size of the data loaded is (2129, 6).

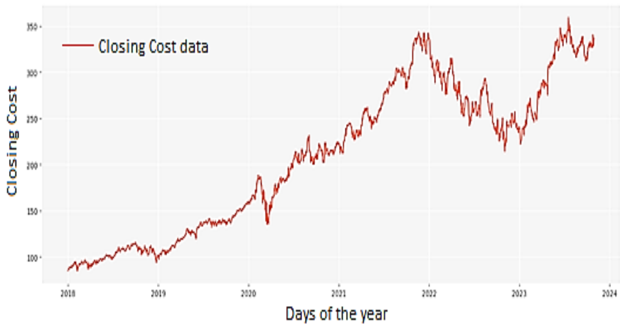
Date	Open	High	Low	Close	Adj Close	Volume
2018-01-01	85.629997	86.050003	85.500000	85.540001	80.008354	18717400.0
2018-01-02	86.129997	86.309998	85.500000	85.949997	80.391846	22483800.0
2018-01-03	86.059998	86.510002	85.970001	86.349998	80.765999	26061400.0
2018-01-04	86.589996	87.660004	86.570000	87.110001	81.476822	21912000.0
2018-01-05	87.660004	88.410004	87.430000	88.190002	82.486984	23407100.0

Fig 5 Example of loaded Microsoft SMP data after 2018 SMP

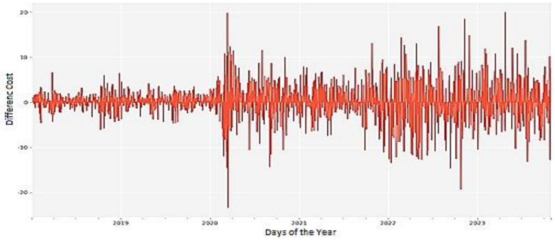
Since the ray time series data is non-stationary in nature thus in this paper it is proposed to make the data stationary the data moving differencing is carried out and a section of data is found to be stationary at first level differencing as shown in Figure 6, thus it is estimated that the value of d lag is set to d=1 for this research.

Among the simplest and most effective ways paper proposed to determine P and Q values using the ARIMA framework is:

- With the data, create a partial autocorrelation graph (PACF). Since P is the threshold point for the PACF, that will assist us in determining the value of P.

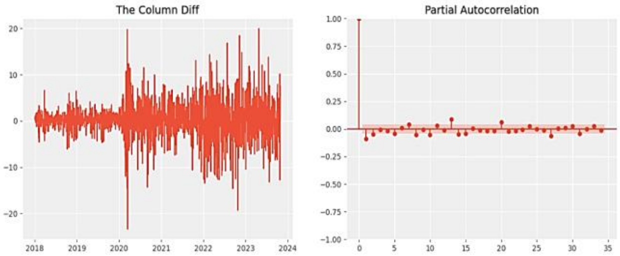


a) raw closing cost data for the last 6 years

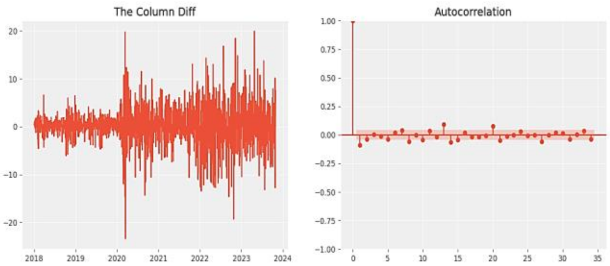


b) results of the first level differencing

Fig 6 results of the data preprocessing implemented for making data stationary



a) Calculate the PACF features for differencing to estimate p lag value



b) Calculate the ACF for q lag

Fig 7 plot of calculated features for differencing to estimate p and q lag values

Stated differently, the order of the AR term is determined by the lags that exceed a certain significance threshold. Take a data autocorrelation graph (ACF). The autocorrelation function (ACF) at lag k, denoted ρk, of a stationary stochastic process, is defined as

$$\rho_k = \frac{\gamma_k}{\gamma_0} \qquad (3)$$

where γ_k *is calcuated as*

$$\gamma_k = cov(yi, yi + k) for\ any\ i. \qquad (4)$$

Since q is the threshold point for the ACF that will assist us in determining the value of q. Paper stated differently the MA term's ordering is determined by the lags that exceed a certain significance ACF threshold

Mathematically the PACF for zero-mean TS stationary data, let ˆxt+h, then for h ≥ 2, denote the regression of TS series defined aa

$$x_t + h\ over\ \{x_t + h - 1,\ x_t + h - 2, \dots, x_t + 1\} \qquad (5)$$

The transformed TS is given as;

$$x\hat{}t + h = \beta_1 x_t + h - 1 + \beta_2 x_t + h - 2 + \dots + \beta_{h-1} x_{t+1} \qquad (6)$$

The PACF of stationary TS $\hat{}t$, represented as φ_{hh} (or φhh), for h = 1, 2, ... is

$$\varphi 11 = corr(xt + 1, xt) = \rho(1) \qquad (7)$$

Based on the aforementioned PACF and ACF graphs as shown in Figure 7 a) and Figure 7 b) respectively, it appears that P=2 and Q=2 have the best values as stated in this proposed work for ARIMA model evaluation.

Result of Stock Data Forecasting

The 100 and 150 sample data is used for demonstrating the data forecasting using the fitted and validated models are used in this paper for representing testing results. The raw open and closing cost data are shown in the Figure 8 below.

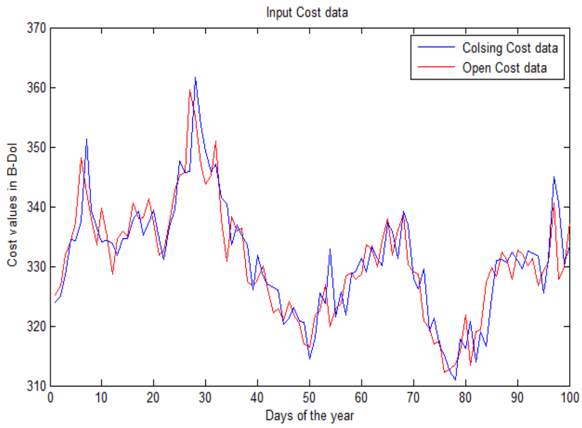


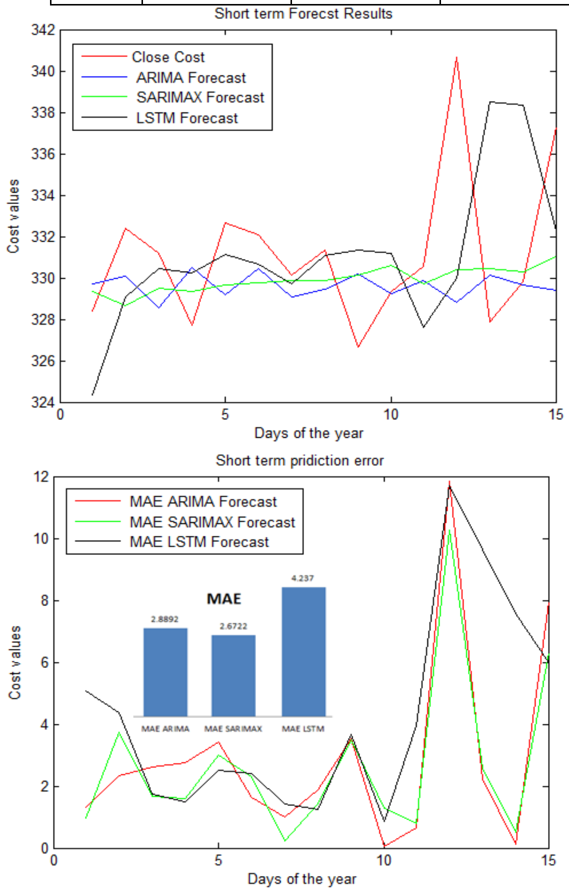
Fig 8 Example of the 100 samples used for Opening and closing price data of Microsoft

Parametric Evaluation

Although the ARIMA model performs exceptionally well, it can be extremely beneficial to integrate seasonality and exogenous variables exclusively. The parametric performance of the short term ARIMA based production model is compared with SARIMAX and LSTM models of prediction for the 15 cost values of closing cost data as shown in the Table 3. It can be observed that ARIMA offers low MSE and is best for short term prediction.

Table 3 Parametric performance comparison for different production models

Error Parameter	ARIMA	SARMAX	LSTM
MSE	9.517423222605261	12.383637904244491	10.862842559814453
MAE	7.2942782308871275	10.001030115447715	9.788872718811035



a) ARIMA/SARIMAX and LSTM prediction

b) ADE and MAE comparison for short term SMP

Fig 9 Results of short-term data prediction with the ARIMA model for Microsoft data of closing cost

Figure 9 displays the outcomes of the ARIMA model's short-term data forecast for Microsoft's closing cost data. The closing cost prediction results are compared for the 15 samples of cost data as shown in the Figure 9 a) for the ARIMA, SARIMAX and LSTM with actual cost. It can be concluded that overall ARIMA performs better. The Figure 9 b) has preseted the error measures as parametric comparison of the performance it can be observed that ARIMMA has lowest performance for the most of the samples. Figure 9 b) plotted the absolute difference error measures. The parameters for the performance measures are defined as

MSE: Mean square error is defied mathematically as the mean of the error square of vectors;

$$MSE = \frac{i}{N}\sum(AC - PC)^2 \quad (8)$$

ADE: Absolute difference error is calculated for each sample of forecast data. As;

$$ADE = {}_{1}^{N}\|(AC - PC)\|_{abs} \quad (9)$$

MAE: Mean Absolute Error, is ciliated mathematically as;

$$MAE = \frac{i}{N}\sum\|(AC - PC)\|_{abs} \quad (10)$$

For the better precise performance of the any prediction model it is essential that the error performance matrices must have the lower values.

7 Result of Long-term LSTM Data Forecasting

In this section the experimental results of the long-term cost prediction for the SMP data are evaluated for the proposed LSTM model with scaled normalization. The architecture of dense SLTM output layers used for proposed LSTM prediction are given in Table 4.

Table 4 the architecture of the dense SLTM output layers

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, 30, 256)	264,192
lstm_1 (LSTM)	(None, 64)	82,176
dense (Dense)	(None, 1)	65

First the basic LSTM model is validated for the random 10lag only and the data prediction is plotted against the raw training data used for model training as shown in the Figure 10.

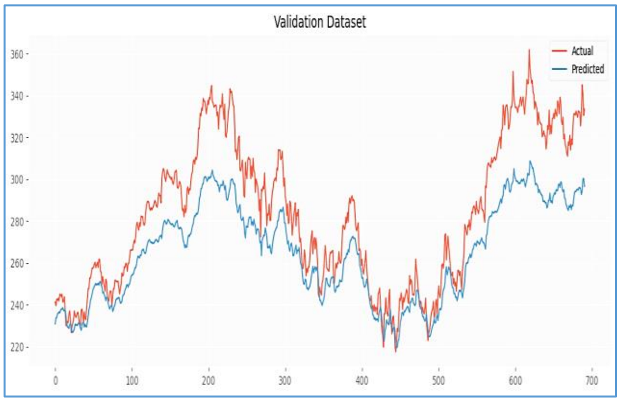
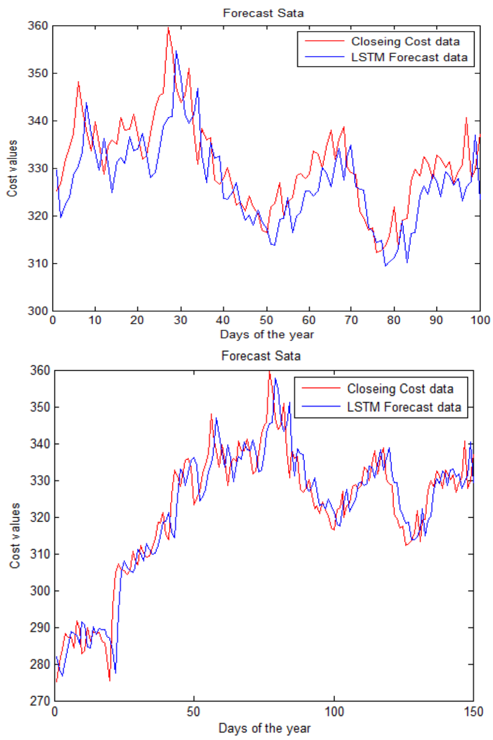


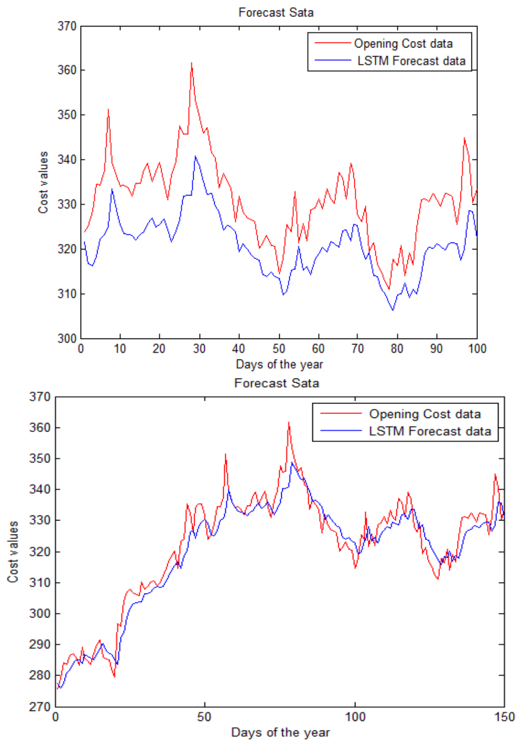
Fig 10 Prediction results with LSTM model validation for closing cost data

It can be observed from the Figure 10 that, there is significant match in the random cross validation of the RNN for LSTM data prediction. But the LSTM parameters are tuned on experimental basis and the dense layer architecture is used further for improving the degree of closeness of the LSTM model. The final prediction results for 100 and 150 samples data for the closing and opening cost are shown in the Figure 11 for LSTM model.



a) Closing cost forecasting for 100 samples

b) Closing cost forecasting for 150 samples



c) Opening cost forecasting for 100 samples

d) Opening cost forecasting for 150 samples

Fig 11 Results of closing and opening cost Forecasting using the proposed long term LSTM model

It is clear from the results of the Figure 11 that the as long is the data sample and as dense the degree of prediction of data is much better. It is concluded that data of 150 samples have much better long-term prediction with the LSTM as compare to 100 samples.

8 Conclusion and Future scope: -

This study offered a comprehensive methodology for the stock market-based prediction. The paper validated the Microsoft database's machine learning-based stock market prediction. This study examined TS prediction as a means of solving the SMP problem. The stock market is a continuously changing sequence of data, which makes cost prediction a challenging subject. This study suggested looking at TS auto regressive models according to the ideal level of differencing in order to improve prediction accuracy. In order to achieve optimal production and a practically stationary model, the lag order is adjusted. The extensive Microsoft stock dataset from the previous 20 years is used to train the prediction methods ARIMA, SARMAX, and LSTM. The ARIMA is suggested for short-term and 150 days for long-term prediction LSTM out performs.

It can be observed that ARIMMA has lowest ADE error performance for the most of the sample and the MAE of 7.2942 is achieved with ARIMA which outperform over SARIMAX and LSTM for short term prediction. But for the long term the LSTM is preferred for the prediction in this research. It is determined that, as compared to 100 samples, the data from 150 samples have substantially better long-term prediction using the LSTM.

In Future it is proposed to compare the performance with the ANN or feed forward networks and the performance is proposed to evaluate for the other IT company dataset also. There is a significant chance of error performance improvement in future, although good accuracy of nearly 93 % is achieved with proposed method.

Acknowledgement

We extend our sincere gratitude to Eklavya University, Damoh, India, for their support in conducting this research titled "Short- and Long-Term Effective Time Series Stock Market Cost Prediction Methods Using Learning Techniques." Additionally, we would like to express our appreciation for the support and guidance provided by our esteemed colleagues at Eklavya University, Damoh, India, namely Sourabh Jain, Dr. Navdeep Kaur Saluja, and Dr. Mohd. Shahnwaz Ansari

References

[1] Pang, X., Zhou, Y., Wang, P. et al. An innovative neural network approach for stock market prediction. Journal of Supercomputing Vol. 76, 2098–2118 (2020). https://doi.org/10.1007/s11227-017-2228-y .

[2] Vaishnavi Gururaj, Shriya V. R, and Dr. Ashwini K., " Stock Market Prediction using Linear Regression and Support Vector Machines," International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 8 pp. 1931-1934 (2019).

[3] Cao, R., Liang, X., Ni, Z. (2012). Stock Price Forecasting with Support Vector Machines Based on Web Financial

Information Sentiment Analysis. In: Zhou, S., Zhang, S., Karypis, G. (eds) Advanced Data Mining and Applications. ADMA 2012. Lecture Notes in Computer Science(), vol 7713. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-35527-1_44

[4] Göçken, M., Özçalıcı, M., Boru, A. et al. Stock price prediction using hybrid soft computing models incorporating parameter tuning and input variable selection. Neural Comput & Applic 31, 577–592 (2019). https://doi.org/10.1007/s00521-017-3089-2

[5] Mehtab, S., Sen, J., Dutta, A. (2021). Stock Price Prediction Using Machine Learning and LSTM-Based Deep Learning Models. In: Thampi, S.M., Piramuthu, S., Li, KC., Berretti, S., Wozniak, M., Singh, D. (eds) Machine Learning and Metaheuristics Algorithms, and Applications. SoMMA 2020. Communications in Computer and Information Science, vol 1366. Springer, Singapore. https://doi.org/10.1007/978-981-16-0419-5_8 .

[6] Kamalakannan J, Indrani Sengupta and Snehaa Chaudhury, "Stock Market Prediction using Time Series Analysis", Computing Communications and Data Engineering Series Volume No 01 Issue No 03 2018 https://ssrn.com/abstract=3168423

[7] Majumder, A., Rahman, M.M., Biswas, A.A., Zulfiker, M.S., Basak, S. (2022). Stock Market Prediction: A Time Series Analysis . In: Somani, A.K., Mundra, A., Doss, R., Bhattacharya, S. (eds) Smart Systems: Innovations in Computing. Smart Innovation, Systems and Technologies, vol 235. Springer, Singapore. https://doi.org/10.1007/978-981-16-2877-1_35

[8] Ariyo, Adebiyi A., Adewumi O. Adewumi, and Charles K. Ayo. "Stock price prediction using the ARIMA model." In 2014 UKSim-AMSS 16th international conference on computer modelling and simulation, pp. 106-112. IEEE, 2014.

[9] Daryl, A. Winata, S. Kumara and D. Suhartono, "Predicting Stock Market Prices using Time Series SARIMA," 2021 1st International Conference on Computer Science and Artificial Intelligence (ICCSAI), Jakarta, Indonesia, 2021, pp. 92-99, doi: 10.1109/ICCSAI53272.2021.9609720.

[10] Alharbi, F.R.; Csala, D. A Seasonal Autoregressive Integrated Moving Average with Exogenous Factors (SARIMAX) Forecasting Model-Based Time Series Approach. Inventions 2022, 7, 94. https://doi.org/10.3390/inventions7040094 .

[11] Sidra Mehtab1 and Jaydip Sen. "A Time Series Analysis-Based Stock Price Prediction Using Machine Learning and Deep Learning Model" International Journal of Business Forecasting and Marketing Intelligence (IJBFMI), Vol 6, No 4, pp. 272 - 335, 2020. https://doi.org/10.1504/IJBFMI.2020.115691

[12] Rouf, N.; Malik, M.B.; Arif, T.; Sharma, S.; Singh, S.; Aich, S.; Kim, H.-C. Stock Market Prediction Using Machine Learning Techniques: A Decade Survey on Methodologies, Recent Developments, and Future Directions. Electronics 2021, 10, 2717. https://doi.org/10.3390/electronics10212717

[13] Shen, J., Shafiq, M.O. Short-term stock market price trend prediction using a comprehensive deep learning system. Journal of Big Data 7, 66 (2020). https://doi.org/10.1186/s40537-020-00333-6

[14] Sakshi Kulshreshtha and Vijayalakshmi A, "An ARIMA-LSTM Hybrid Model for Stock Market Prediction Using Live Data", Journal of Engineering Science and Technology Review 13 (4) (2020) 117 – 123 http://www.jestr.org/

[15] Shakir Khan, Hela Alghulaiakh, "ARIMA Model for Accurate Time Series Stocks Forecasting", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 11, No. 7, 2020, http://www.ijacsa.thesai.org/

[16] Aminimehr, A., Raoofi, A. ,et al. A Comprehensive Study of Market Prediction from Efficient Market Hypothesis up to Late Intelligent Market Prediction Approaches. Comput Econ 60, 781–815 (2022). https://doi.org/10.1007/s10614-022-10283-1

[17] Sidharth Tiwary, Pramod K. Mishra, "Time-series Forecasting of Stock Prices using ARIMA: A Case Study of TESLA and NIO" Proceedings of the 2nd Indian International Conference on Industrial Engineering and Operations Management Warangal, Telangana, India, August 16-18, 2022 xxx

[18] Jimmy Ming-Tai Wu Zhongcui Li Norbert Herencsar Bay Vo, Jerry Chun-Wei Lin "A graph-based CNN-LSTM stock price prediction algorithm with leading indicators" Multimedia Systems (2023) 29:1751–1770 https://doi.org/10.1007/s00530-021-00758-w

[19] Kelum Gajamannagea, , Yonggi Park "Real-time Forecasting of Time Series in Financial Markets Using Sequentially Trained Many-to-one LSTMs" arXiv:2205.04678v1 [cs.LG] 10 May 2022

[20] Chance Robinson, Nnenna Okpara, Michael Burgess, Faizan Javed "Stock Forecasts with LSTM and Web Sentiment" SMU Data Science Review, Vol. 6 [2022], No. 2, Art. 10 https://scholar.smu.edu/datasciencereview/vol6/iss2/10

[21] Salman Ahmed , Saeed-Ul Hassan , Naif Radi Aljohani , Raheel Nawaz "FLF-LSTM: A novel prediction system using Forex Loss Function" DOI: https://doi.org/10.1016/j.asoc.2020.106780

[22] Tawum Juvert Mbah & Haiwang Ye & Jianhua Zhang & Mei Long "Using LSTM and ARIMA to Simulate and Predict Limestone Price Variations" Received: 10 April 2020 /Accepted: 16 November 2020 Published online: 6 January 2021 https://doi.org/10.1007/s42461-020-00362-y

[23] Faraz Ahmad , Pawan Singh "Stock Price Prediction using ML and LSTM based Deep Learning models" Journal of Informatics Electrical and Electronics Engineering, 2022, Vol. 03, Iss. 01, S. No. 008, pp. 1-13 ISSN (Online): 2582-7006

[24] Hum Nath Bhandari, Binod Rimal, Nawa Raj Pokhrel, Ramchandra Rimal, Keshab R. Dahal, Rajendra K.C. Khatri, "Predicting stock market index using LSTM", Machine Learning with Applications, Volume 9, 2022, 100320, https://doi.org/10.1016/j.mlwa.2022.100320 .